Weekly Report(July.15.2019-July.21.2019)

Shanghai Jiao Tong University
1747836307@qq.com
Shaopu Rui

Abstract

This week I continue to learn VGG net.

1 Problems left last week

- The accuracy got 80.19% after 100 epochs when setting learning rate to be 0.01 and batch size to be 128.
- The smaller the learning rate, the worse the accuracy.
- Learning rate is actually the problem that causes the result not so good. But after several experiments, Adadelta still didn't get better performances than SGD.

Table 1: batch_size=128, optimizer=Adadelta

Learning Rate	Accuracy
0.01	80.19%
0.003	75.47%
0.001	73.54%
0.0003	67.21%
0.0001	54.53%

2 VGG

- At the very beginning, I made a stupid mistake that I started to train the VGG model imported from pytorch but not pretrained. I forgot to add the parameter **pretrained=true**. Obviously it performs not good.

Table 2: momentum=0.9, optimizer=SGD

Learning Rate	Batch Size	Accuracy
0.0003	8	91.10%
0.0003	16	90.26%
0.0003	32	89.34%
0.0001	8	90.19%
0.0001	16	87.79%
0.0001	32	82.30%

- Then I started to train the whole pretrained model with 2 epochs. The result are shown in table 2. It gives the best acuuracy of 91.10%.

  - The learning rate can't be too large like 0.01 otherwise loss will explode.

- I also tried to only train the fully connected layers and only train the last full connected layer. The results are shown in table 3 and table 4. They didn't perform so well than I thought.
  - A possible reason is all my trainings are based on a pretrained model. When I train all layers in several epochs, it's more likely to perform to better in cifar10 dataset, which is smaller than ImageNet.

Table 3: momentum=0.9, epochs=2

Learning Rate	Batch Size	Accuracy
0.003	8	81.59%
0.003	16	84.02%
0.003	32	83.98%
0.001	8	85.05%
0.001	16	85.77%
0.001	32	85.43%
0.0003	8	85.73%
0.0003	16	85.80%
0.0003	32	84.28%
0.0001	8	85.05%
0.0001	16	82.84%
0.0001	32	79.96%

Table 4: momentum=0.9, epochs=2

Learning Rate	Batch Size	Accuracy
0.003	8	81.16%
0.003	16	83.05%
0.003	32	82.98%
0.001	8	83.22%
0.001	16	83.36%
0.001	32	82.92%
0.0003	8	83.35%
0.0003	16	82.27%
0.0003	32	81.53%
0.0001	8	90.19%
0.0001	16	87.79%
0.0001	32	82.30%

3 Plan

- Fine tuning neural network needs more patience, I will start to use pre-trained ResNets to build model and fine tune the model with cifar10. Time is limited, I will do as more as possible.